DOCUMENT RESUME

ED 371 622 FL 022 269

AUTHOR Pavlou, Pavlos Y.; Rasi, Sylvia B.
TITLE Considerations in Writing Item Prompts for Simulated Oral Proficiency Interview (SOPI).
PUB DATE Mar 94
NOTE 16p.; Paper presented at the Annual Language Testing Research Colloquium (Washington, DC, March 1994).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Bilingual Education; *Language Proficiency; Language Teachers; *Language Tests; *Oral Language; Second Language Instruction; Spanish; Teacher Certification; Teacher Education; Test Construction; *Test Content; *Test Items; Test Validity
IDENTIFIERS Oral Proficiency Testing; *Simulated Oral Proficiency Interview; Texas

ABSTRACT

This study addresses the validation of an oral proficiency test by examining the language that is obtained from the items presented in the test. The study focused on the language elicited by a form of the Simulated Oral Proficiency Interview (SOPI), the Spanish version of the Texas Oral Proficiency Test (TOPT). The TOPT is used to measure the oral proficiency of foreign language and bilingual education teachers seeking certification in Texas. The result of an analysis of the responses of a number of examinees representing a wide range of overall proficiency levels was the development of a typology of SOPI characteristics. The broad categories of the typology are speaker, topic, audience, and setting. Inter-item relationships and grammatical features of the test were also examined. An appendix outlines the typology of prompt characteristics. (MDM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 371 622

CONSIDERATIONS IN WRITING ITEM PROMPTS FOR SIMULATED ORAL PROFICIENCY INTERVIEW (SOPI)

PAVLOS Y. PAVLOU and SYLVIA B. RASI

Georgetown University, Washington D.C.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Pavlou

Pavlou

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Paper presented at annual Language Testing Research Colloquium
Washington, D.C., March 1994

FL 022267

BEST COPY AVAILABLE

Consideration in Writing Prompts for Simulated Oral Proficiency Interviews (SOPI)

Pavlos Y. Pavlou and Sylvia B. Rasi[1]

INTRODUCTION

This paper addresses the validation of an oral proficiency test by examination of the language that is obtained from the items presented in the test. As Shohamy, Shmueli and Gordon (1991) suggest, correlation validation studies of language tests on their own are insufficient for their evaluation. "There is a need to take a broader and deeper look at the language tests from more expanded perspectives - its elicitation tasks, the specific language that is obtained as a result of these tasks and the specific strategies that test takers employ in taking tests." (p. 2) As Shohamy et al. suggest, "Even in situations when it is not possible to carry out a large scale construct validation study there are still many other ways to obtain evidence for validity tests" for multiple perspectives are necessary in this process (ibid.). This type of approach is supported by a number of researchers concerned with the validity of oral proficiency tests (e.g., Shohamy, 1988; Grotjahn, 1986; and Vollmer, 1981). Our study of the Simulated Oral Proficiency Interview (SOPI) represents a step in this direction.

Continuing attention to communicative competence in teaching and speaking of foreign languages within the past fifteen years (Canale and Swain, 1980; Omaggio, 1986) has generated a corresponding increase in activity in the development of authentic communicative tests of oral proficiency. One of the methods of

testing speaking proficiency is the Oral Proficiency Interview (OPI), a face-to-face interview based on the Speaking Proficiency Guidelines developed in the 1970s by U.S. government agencies and adopted by the American Council for the Teaching of Foreign Languages (ACTFL).

The SOPI has followed the OPI's lead but instead employs a semi-direct method, employing a cassette tape, player, and test booklet to measure speaking proficiency (Clark, 1988; Stansfield and Kenyon, 1988; Stansfield et al., 1989; Stansfield 1989). Semi-direct tests of English as a Foreign Language were developed in 1982 by the Educational Testing Service; however, the SOPI differs from these in that the SOPI is ACTFL-based and visual as well as aural stimuli are employed. While Shohamy, Gordon, Kenyon, and Stansfield (1988) found the OPI and the SOPI to have high correlations with one another (above r = .90), the development of the SOPI represents an important advance in the testing of oral proficiency in that it provides an efficient, accurate and economical method which provides a uniform testing situation. Thus, the SOPI merits attention and analysis. The present paper investigates issues involved in writing items for the SOPI and makes suggestions regarding considerations which should be incorporated into future SOPIs.

The levels of oral proficiency established by the ACTFL Guidelines range from Novice to Superior. The SOPI is designed to test the Intermediate to Superior levels. Among other features, the Intermediate level of proficiency is characterized by the

ability to ask and answer questions, and sustain a simple conversation on everyday topics. Discourse at this level is generally limited to the sentence level. The Advanced level is characterized by the ability to narrate, describe, and explain in past, present, and future time about topics of current and personal interest. At this level, paragraph-length discourse is evident. The Superior level rating is assigned to individuals who are able to participate in formal and informal conversations on practical, social, and professional topics. Those performing at the Superior level are able to support opinions and hypothesize.

All items of the SOPI consist of an explanation of the task in English followed by an utterance in the target language. The target language utterance, or prompt[1], indicates that the examinee may begin his or her response. In addition, the target language prompt allows the examinee to hear some of the target language prior to beginning the response. It may provide vocabulary and at times serves to remind the examinee of the context. Other than the target language itself, the prompt does not provide instructions or information that is not in the English language explanation. In other words, this arrangement makes it possible for the examinee to respond to the item without having to necessarily understand the target language utterance provided. This method ensures that the test is designed to measure the examinee's speaking performance without measuring listening skills as well.

The present paper is based on an analysis of items designed for the Spanish version of the Texas Oral Proficiency Test (TOPT),

a form of the SOPI developed by the Center for Applied Linguistics for the State of Texas (Stansfield and Kenyon, 1991). Since 1990, the French and Spanish TOPTs have been used to measure the oral proficiency of foreign language and bilingual education teachers seeking certification in the State of Texas. A rating at the Advanced level of speaking proficiency is required to pass the test. Several other SOPIs, targeted at a more general audience (e.g., Chinese, Portuguese, Hebrew, Hausa, Indonesian) as well as high school and college level students (e.g., Japanese, Spanish) have since been developed, with several more in the process of development (e.g., French, German).

This paper discusses issues in writing prompts for Simulated Oral Proficiency Interviews (SOPI) and attempts to devise a typology of characteristics of these prompts. The relationship among the various prompts is also investigated.

OBJECTIVE

This study was primarily undertaken in order to gain a better understanding of which item characteristics assisted examinees to achieve the minimum level of proficiency required by a particular item. This is an important objective given the test's stated purpose of providing an accurate picture of the examinee's current ability to speak the target language in a variety of circumstances and on a variety of topics. A secondary objective was to understand the relationships between the items in the entire test. In addition, we studied which grammatical structures were present in examinees' responses and evaluated whether they coincided with

the goals of the item. Duplication of grammatical structures found in examinee responses was also considered across items. We believe that when any new method or instrument is developed for measuring language proficiency this type of "after-the-fact" analysis complements the extensive research conducted prior to the development of the test, provides highly valuable information for future, improved versions of the tests, and informs researchers developing alternative measures.

METHOD

Items from the Spanish TOPT (Texas Oral Proficiency Test) and responses elicited from a selected number of examinees representing a wide range of overall proficiency levels were analyzed with the above objectives in mind. Through repeated listening and study of transcripts we attempted to identify the part of the item which functioned as a trigger for specific segments of the examinee's response. We first considered how similar item contents elicited similar responses and then reformulated the items into more general situations and sets of instructions.

The result of this analysis was the development of a Typology of SOPI Item Characteristics (Appendix A) which we hope will be useful to future developers of semi-direct tests of oral proficiency.

The broad categories of the typology are: speaker[3], topic, audience, and setting. Different configurations within these categories result in items that vary in difficulty. While lower level prompts do necessarily reflect all four categories, these

categories are found in all Advanced and Superior level items. An example of an item together with its analysis in terms of these categories follows:

> A group of high school students has arrived from Colombia to spend a summer session at a community college in Texas. You have been asked to give a brief talk as part of their orientation on two or three recent events in Texas that you feel are important. After your talk is introduced, brief the group on these events.

This item was reformulated in terms of the four categories listed above with the following result:

TABLE 1. SAMPLE ITEM ANALYSIS.

SPEAKER: Speaker's point of view: Personal

TOPIC: Familiarity with topic: Speaker probably has knowledge of current events

AUDIENCE:
Size: Greater than two persons

Speaker-audience relationship:
Age: Speaker is older than audience
Status: Speaker is of higher status than audience
Audience knowledge: No expert knowledge required

SETTING: In the speaker's country/culture

INTER-ITEM RELATIONSHIPS

Since the content of the responses could not be solely accounted for by their respective items, an analysis of the relationship between the items was also undertaken. Regarding the relationship between items, we discovered that substantial shifts in the speaker's assigned roles between items resulted in "tailoring problems", i.e., difficulty in responding to the item in

an appropriate manner, taking into consideration the task and audience. This is noted by Shohamy et al. (1991) who note that "in spite of the fact that the designers of the SOPI tried to introduce an hypothetical interlocutor it seems that the test takers do not really consider that interlocutor that much. The test taker is more concerned about the completion of the task as dictated by the question on the tape" (p. 14)

An example of this discontinuity is reflected by Situations 4 and 5 of the Spanish TOPT:

Situation 4

> You are an exchange teacher at a school in Monterrey, Mexico. You have been asked to give a workshop for teachers on increasing student involvement in class activity. After your talk is introduced, **first, convey your pleasure at being invited to give the talk and then begin to discuss ways to increase student participation in the class.** Note that you are not expected to give a complete talk in the time allotted; you only need to begin your talk and continue until your time is up.

The above item requires the speaker to assume an authoritative and formal role. Speakers responded to the requirements of the task which elicited formal language. Situation 5, which immediately follows Situation 4 on the TOPT, represents a shift in the speaker's role.

Situation 5

> You are an exchange student at Veracruz, Mexico. A colleague and friend of yours, Ana Marcos, who teaches English, is considering participating in a year-long teacher exchange program in the U.S. She has never been to the U.S. before. One day she comes to you asking for your advice on whether or not she should participate. After she asks her question, **advise her, from your point of view, on what she should do.**

Upon analysis, the above item is seen to require the speaker to assume a role of an equal in a non-authoritative, informal situation. However, analysis of the taped responses revealed two cases where speakers used language which was more formal than would be necessary in such a situation. In responding to Situation 5, the speakers seemed to remain in the formal "mode" required by Situation 4. We believe these and similar inappropriate responses can be attributed to the discontinuity in the roles set up by consecutive items.

Elicitation of Grammatical Features

In analyzing the examinee responses it became apparent that certain grammatical features such as verb tense were being elicited more than once during the course of the test, resulting in a less that efficient use of item potential to elicit a wide variety of ratable speech.

An example of this problem is found in Picture #2, whose task is "Describe a place or activities". As expected, the item elicited many present tense forms. The following task (picture #3) asks the examinee to "Narrate in Present Time" and was also aimed at eliciting this grammatical form.

A second overlapping feature was found in the items listed below:

Item	Task	Grammatical Feature
Picture 1	Give directions	Imperative forms
Picture 3	Narrate at present time	Present tense forms
Picture 4	Narrate in past time	Past tense forms
Picture 5	Narrate in future time	Future tense forms

The grammatical features aimed at above do not overlap. However, at the same time all prompts tend to elicit temporal/sequential conjunctions, for example: primero, segundo, entonces, finalmente (first, second, than, finally). While in general it must be recognized that some grammatical duplication of this nature is inevitable, we believe with appropriate planning based on analyses such as this one, greater efficiency in SOPI item construction will be facilitated.

Conclusion

While far from exhaustive, our analysis of SOPI items indicated that there are a number of important considerations which must be taken into account in the item-writing process for these types of tests. Such characteristics include audience size, social distance between the speaker and the audience, and the degree of formality of the situation. Moreover, our analysis revealed that an inability to adequately define one of these item characteristics for a given item in a manner that is understood by the examinee may result in an unexpected and undesirable examinee response. In the future, we suggest that items be constructed and analyzed taking these preliminary categories into account.

In terms of inter-item considerations, it was found that the consistency of the roles which items assign to the speaker is a factor which affects the speaker's adequate tailoring of the response. Since SOPIs do not follow a linear model of difficulty (e.g., all Intermediate items followed by all Advanced items, etc. following the OPI "level probes") it may be somewhat difficult to

maintain consistency in the roles to be assumed by the speakers. However, we suggest that use of the "natural" breaks in the test (e.g., breaks between the Pictures, Topics, and Situations sections) together with greater inter-sectional planning could bring about the best possible situation in eliciting appropriate responses.

This preliminary analysis of item characteristics together with the Typology of SOPI Item Characteristics (Appendix A) can be used and expanded by prompt writers in the future to ensure that a given prompt triggers the response for which the item is targeted and contributes possible to the overall SOPI in the most efficient way. We encourage researchers and test developers to work in this direction.

APPENDIX A : TYPOLOGY OF PROMPT CHARACTERISTICS

Prompt Characteristics	Minimum Level Required
Speaker	
I. Speaker's point of view	
1. Speaker speaks from his/her personal point of	I
2. Speaker speaks from a group's point of view	
a. Which s/he agrees with	I
b. Which s/he does not agree with	
Topic	
Speaker-Topic relationship	
I. Familiarity/expertise of speaker with topic	
1. Topic from everyday life that	
a. Speaker has most probably experienced	I
b. Speaker has not directly experienced but s/he can easily imagine the topic and the situation (no hypothesizing required).	I
c. Speaker has no experience with the topic or the situation (needs to hypothesize).	A
2. Topic requires more than general knowledge on the part of the speaker	
a. Speaker has this knowledge either through interest or education	A
3. Speaker has no knowledge of the topic either through interest or education	A
II. Nature of Topic	
1. Degree of interest	
a. Topic is interesting to the audience. Therefore, less effort is required on the part of the speaker to attract and maintain the interest of the audience.	
b. Topic is not very interesting to the audience. Therefore, more effort is required on the part of the speaker to attract and maintain interest.	S
2. Effect on audience	
a. Topic is pleasant to the audience.	I
b. Topic is not pleasant to the audience. Therefore, there is more burden on the speaker not to insult or offend anyone.	A

3. Speaker transferring what other people believe about a certain topic (someone else's opinion).
   a. Speaker agrees with the group's point of view. A
   b. Speaker distances himself from the opinion he is expressing. A+

Audience

I. Size
   a. Audience consists of one person I
   b. Audience consists of two persons I
   c. Audience consists of more than one person I

II. Speaker-audience relationship
   a. Age
      S younger than A
      S has same age with A
      S older than A
   b. Status
      S superior to A
      S has equal status with B
      S is inferior to A
   c. Degree of intimacy
      friend, stranger, professional peers

IV. Knowledge assumed from audience
   a. Topic requires no 'expert' knowledge on the part of the audience.
   b. Topic requires expert knowledge on the part of the audience.
      1. Audience has this knowledge. A

      2. Audience does not have this knowledge. Therefore, speaker is responsible to explain things the audience is not familiar with. S

Where the conversation takes place

   a. In the speaker's country (culture)
   b. In the foreign country (culture)

REFERENCES

Canale, M. and Swain, M. 1980. Theoretical basis of communicative approaches to second language teaching and testing. *Applied Linguistics, 1,* 1-47.

Clark, J.L.D. 1975. Theoretical and technical considerations in oral proficiency testing. In R.L. Jones and B. Spolsky (eds.), *Testing Language Proficiency* (pp. 10-28) Arlington, VA: Center for Applied Linguistics. Educational Testing Service. 1982. Oral proficiency testing manual. Princeton, NJ: ETS.

Educational Testing Service. 1982. *Oral Proficiency Testing manual.* Princeton, NJ: ETS

Grotjahn, R. 1986. Test validation and cognitive psychology. *Language Testing, 3,* 159-185. Omaggio, A. 1988. Teaching Language in Context. Boston, MA: Heinle & Heinle.

Shohamy, E. 1988. A proposed framework for testing the oral language of second/foreign language learners. *Studies in Second Language Acquisition, 10,* 165-79.

Shohamy, E., Gordon, C., Kenyon, D., and Stansfield, C. 1989. The development and validation of a semi-direct test for assessing oral proficiency in Hebrew. *Bulletin of Higher Hebrew Education.* Vol. 4.

Shohamy, E., Scmueli, D., and Gordon, C. 1991. The validity of concurrent validity of a direct vs semi direct test of oral Proficiency. Paper presented at the 13th Language Testing Research Colloquium. Princeton, March.

Stansfield, C., and Kenyon, D.M. 1988. *Development of the Portuguese Speaking Test.* Year one project report on the development of semi-direct tests of oral proficiency in Hausa, Hebrew, Indonesian and Portuguese. Alexandria, VA: ERIC Reproduction Service. ED 296 586.

Stansfield, C. and Kenyon, D.M. 1988. *Development of Semi-Direct Tests of Oral Proficiency in Hausa, Hebrew, Indonesian and Portuguese.* Washington, DC: Center for Applied Linguistics.

Vollmer, H. 1981. Why are we interested in "General language Proficiency?" In Klein-Brawley, C. and Sevenson, D.K. (eds.), *Practice and Problems in Language Testing 1,* Bern: Lang.

NOTES

[1]The name of the authors appear in alphabetical order

[2]There has been some variation in the use of the term prompt. In an attempt to provide greater clarity, we have employed the term item to indicate the English language explanation together with the target-language utterance. We have designated the term prompt to refer exclusively to the target-language utterance following the English language explanation. (In the past, the term prompt has referred to our item.)

[3]The term speaker is used interchangeably with examinee.